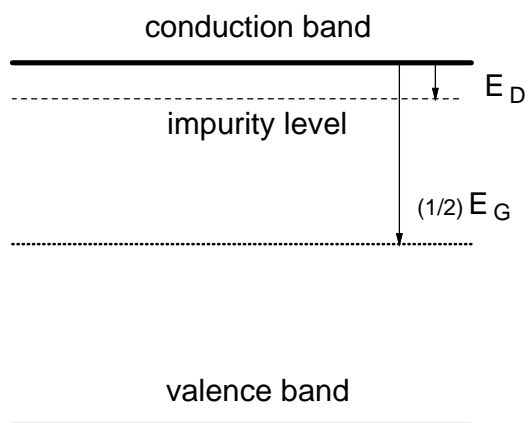
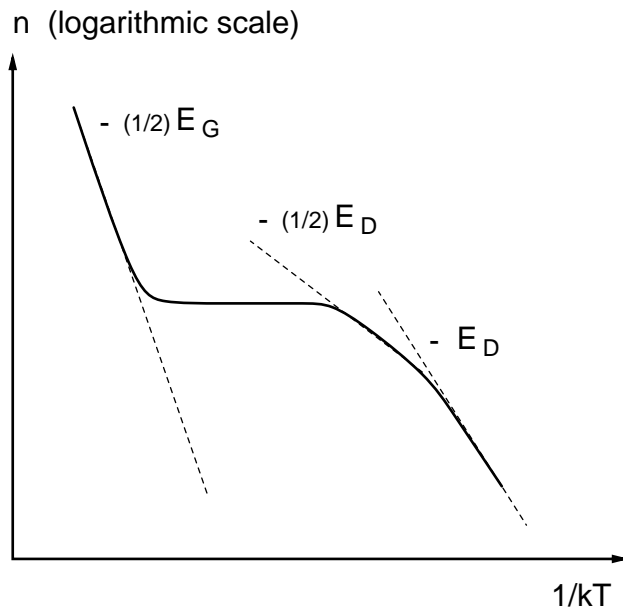


Basic Semiconductor Physics, wrap-up

Equilibrium Statistics



Equation that determines the Fermi level E_F is obtained from the neutrality condition:

$$n + n_D = N_D - N_A, \quad (3.1)$$

where N_D and N_A are the concentrations of donors and acceptors, $N_D > N_A$, n is the band electron concentration,

$$n = N_C F_{1/2}(E_F/kT) \approx N_C e^{E_F/kT}, \quad (3.2)$$

and n_D is the concentration of electrons residing on donors (donor energy, counted from the conduction band is $E_D < 0$):

$$n_D = \frac{N_D}{\frac{1}{2} e^{(E_D - E_F)/kT} + 1}. \quad (3.3)$$

Denoting $E_F \equiv e\mu$, and $\beta \equiv 1/kT$ or $\beta \equiv e/kT$, depending on whether β multiplies an energy or a potential, we rewrite Eq. (3.1) in two equivalent forms:

$$N_C e^{\beta\mu} + \frac{N_D}{\frac{1}{2} e^{\beta(E_D - \mu)} + 1} = N_D - N_A, \quad (3.4)$$

$$n^2 + (N_A + N_C^*) n - (N_D - N_A) N_C^* = 0, \quad (3.5)$$

where
$$N_C^* \equiv \frac{1}{2} N_C e^{\beta E_D} \quad (3.6)$$

Solution of Eq. (3.5) is of the form

$$n = -\frac{N_A + N_C^*}{2} + \frac{1}{2} \left[(N_A + N_C^*)^2 + 4(N_D - N_A)N_C^* \right]^{1/2}. \quad (3.7)$$

$$N_A^2/N_D \ll N_C^* \ll N_D \quad n = \sqrt{(N_D - N_A)N_C^*} = \sqrt{\frac{1}{2}N_D N_C} e^{\frac{1}{2}\beta E_D} \quad (3.8)$$

$$N_C^* \ll N_A^2/N_D \quad n = \frac{(N_D - N_A)N_C^*}{N_A} = \frac{N_D N_C}{2N_A} e^{\beta E_D} \quad (3.9)$$

The range (3.8) occurs at not too low temperatures when $N_A \ll n \ll N_D$ (the existence of this range presupposes $N_A \ll N_D$). The steeper temperature dependence (3.9) occurs at lower temperatures, when $n \ll N_A$. In an artificial example of $N_A=0$ (*totally uncompensated semiconductor*) the range (3.9) does not exist. In both of these low- T ranges the Fermi level lies *above* the donor level:

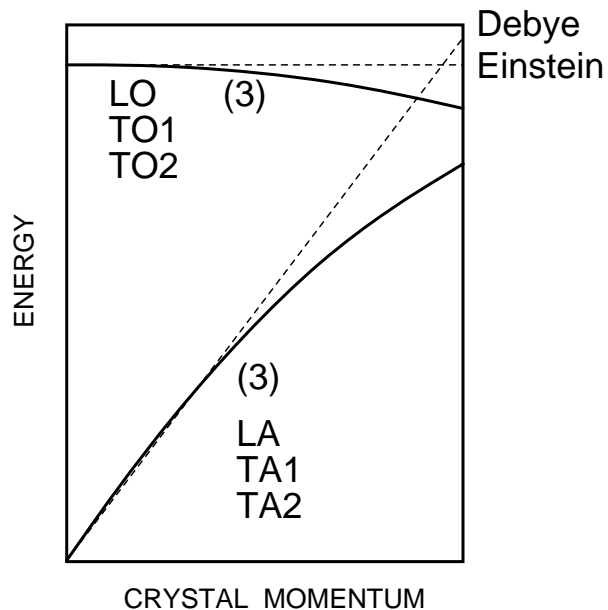
$$E_F = \frac{1}{2} E_D + \frac{1}{2} kT \ln (N_D/2N_C) > E_D \quad N_A \ll n \ll N_D \quad (3.10)$$

$$E_F = E_D + kT \ln [(N_D - (N_A)/2N_A)] > E_D \quad n \ll N_A \quad (3.11)$$

For $N_A=0$ (unphysical situation often treated in books) $E_F(T \rightarrow 0) = \frac{1}{2} E_D$; for any realistic compensation $E_F(T \rightarrow 0) = E_D$. The slope of $E_F(T)$ is positive at $T=0$ (except if $N_A > N_D/3$).

At higher temperatures, E_F goes below E_D . When $E_D - E_F \gg kT$ (saturation regime), most donors are ionized and all available excess electrons are in the conduction band.

Lattice vibrations



Debye model describes Acoustic phonons pretty well.
Einstein model is more appropriate for Optic phonons.
Both models can be used simultaneously.

With the help of these models, one can successfully describe even *anisotropic* properties of the phonon field, see e.g.,

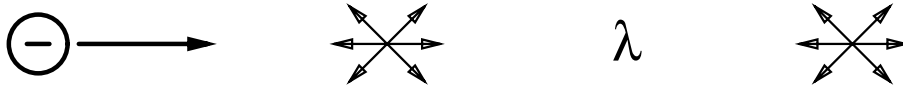
S. Luryi and J. Van Kranendonk, "Elastic constants and anisotropic pair correlations in solid hydrogen and deuterium", *Can. J. Phys.* **57**, pp. 136-146 (1979).

Optic phonons which strongly interact with electrons are essentially *monochromatic*. This can have amazing consequences, see, e.g. A. A. Grinberg and S. Luryi (1990)

"Nonstationary quasiperiodic energy distribution of an electron gas upon ultrafast thermal excitation", *Phys. Rev. Lett.* **65**, pp. 1251-1254 (1990).

"Fine structure in the energy dependence of current density and oscillations in the current-voltage characteristics of tunnel junctions", *Phys. Rev. B* **42**, pp. 1705-1712 (1990).

Transport



$$\frac{d\mathbf{v}}{dt} = \mathbf{a} - b\mathbf{v}$$

$$\begin{array}{ll} \mathbf{v} = \mathbf{a} t & t \ll \tau \\ \mathbf{v} = \mathbf{a}/b & t \gg \tau \end{array}$$

τ : velocity (momentum) relaxation time

Loosely speaking: τ is average time $\langle t \rangle$ between collisions

After "collision" (at $t = 0$) take $\langle \mathbf{v} \rangle = 0$

$$\mathbf{v}(t) = \frac{e}{m} t \mathbf{F}$$

$$\langle \mathbf{v} \rangle = \frac{e}{m} \tau \mathbf{F} \equiv \mu \mathbf{F}$$

$$\mu = \frac{e \tau}{m} \quad \text{Drude formula}$$

Drift Diffusion Equation

$$J = e (n \mu \mathbf{F} + D \nabla n)$$

Mobility and diffusion coefficients are related (Einstein)

$$eD = \mu kT$$

☞ from the consideration that **in equilibrium** one must have $J \rightarrow 0$.

$$e \mathbf{F} = \nabla E_C$$

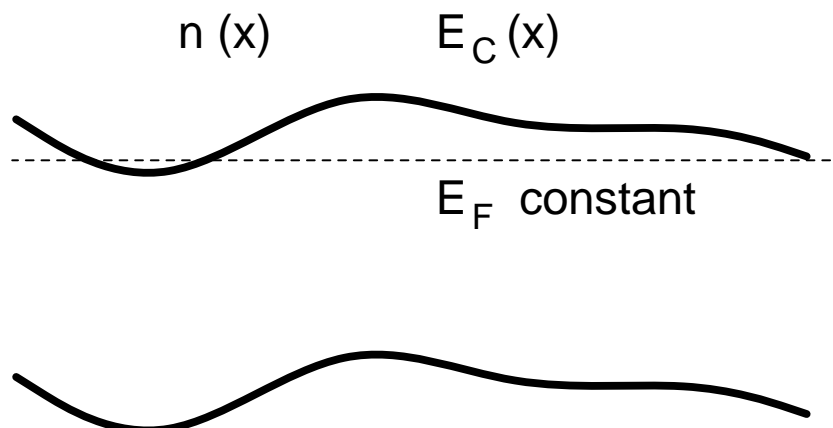
$$\nabla n = -\frac{n}{kT} \nabla E_C$$

☞ the latter equation is obtained by differentiating

$$n = N_C e^{(E_F - E_C)/kT}$$

Degenerate version of Einstein relation:

$$eD \frac{d(\log n)}{dE_F} = \mu$$



Imref

In the presence of current flow (non-equilibrium) the concept of Fermi level is not defined.

However, let us **define**

$$n = N_C e^{(E_F - E_C)/kT}$$

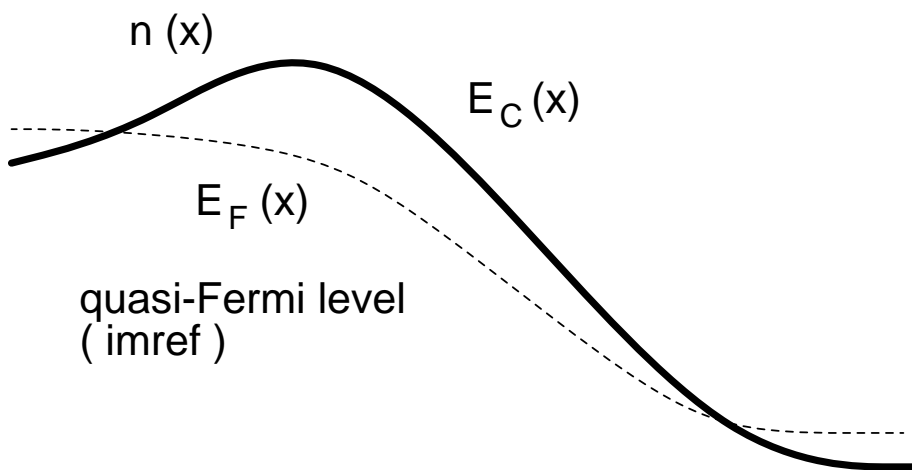
where

$$\begin{aligned} n &= n(x) \\ E_C &= E_C(x) \\ E_F &= E_F(x) \end{aligned}$$



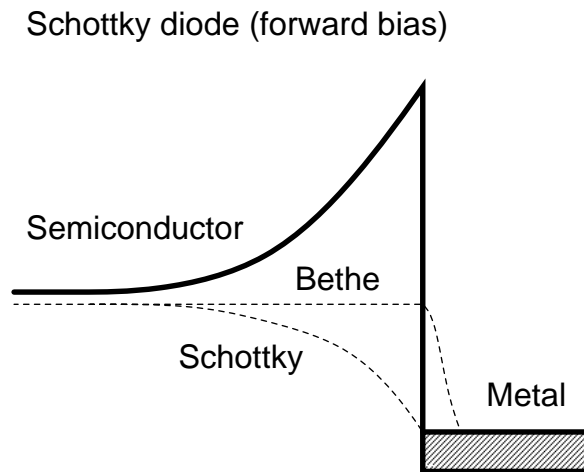
The drift diffusion equation then reduces to

$$J = e (n \mu \mathbf{F} + D \nabla n) \equiv e n \mu \nabla E_F$$



Example: use of imref to distinguish "pictorially" different transport mechanisms.

Thermionic (Bethe) versus diffusion (Schottky) mechanism of conduction in Schottky diodes.



☞ Imref varies little where the net current is much smaller than either the diffusion or the drift components. These regions are "approximately in equilibrium".

Conduction Laws by dimensional analysis§

In the CGS system the conductivity has the units of a velocity.† Taking this velocity to be an effective carrier velocity v , we can write a generic expression for the diode current in the form

$$I \propto \frac{\varepsilon}{4\pi} v V \frac{A}{L^2}, \quad (\text{bulk})$$

$$I \propto \frac{\varepsilon}{4\pi} v V \frac{W}{L}. \quad (\text{film})$$

where L is the length, $A = D \cdot W$ the cross-sectional area, and W the width of the diode. The relative permittivity $\varepsilon \equiv \varepsilon/\varepsilon_0$ of the material enters because it scales the space-charge potential in Poisson's equation.

The actual current-voltage dependence (up to a numerical coefficient) can be "derived" from the above equation – whenever the conduction process involves a dominant transport mechanism, which provides a unique scaling relationship between v and V .

Thus, for free electron motion, the velocity scales as $v^2 \propto (e/m) V$ and one obtains laws appropriate for ballistic transport, e.g. for the bulk case Child's law of vacuum electronics:

$$I = \zeta \frac{\varepsilon}{4\pi} \left[\frac{e}{m} \right]^{1/2} V^{3/2} \frac{A}{L^2} \quad \text{where} \quad \zeta = \frac{4\sqrt{2}}{9} \quad [\text{Child}]$$

For the case when electron velocity is saturated, take $v = v_S$.

For the case of constant mobility μ , the velocity scales as $v \propto \mu V/L$, which leads to the following expressions:

$$I = \zeta_3 \frac{\varepsilon}{4\pi} \frac{\mu V^2}{L^3} A, \quad (\text{bulk}) \quad \text{where} \quad \zeta_3 = \frac{9}{8} \quad [\text{Mott-Gurney}]$$

$$I = \zeta_2 \frac{\varepsilon}{4\pi} \frac{\mu V^2}{L^2} W. \quad (\text{film})$$

§ S. Luryi, "Device building blocks", Chap. 2 in *High-Speed Semiconductor Devices*, ed. by S. M. Sze, Wiley Interscience (1990) pp. 57-136.

† Connection to the international units is obtained by replacing the dimensionless factor $\frac{\varepsilon}{4\pi}$ with ε in farads per meter.

Nonequilibrium electron ensembles.†

The term “hot electrons” purports a non-equilibrium ensemble of high-energy carriers. It is often possible to pump external energy (e.g., by shining light or applying an electric field) directly into the system of carriers. If the power input into the electronic system exceeds the rate of energy loss by that system to the lattice, then the carriers “heat up” and their velocity distribution $f(\mathbf{v})$ deviates significantly from the equilibrium Maxwellian form.§ In general, the time-dependent distribution function $f(t, \mathbf{r}, \mathbf{v})$ can be determined by solving the Boltzmann transport equation,

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + \mathbf{a} \cdot \frac{\partial f}{\partial \mathbf{v}} = \left[\frac{\partial f}{\partial t} \right]_{\text{coll}},$$

where \mathbf{a} is the acceleration and the collision integral in the right-hand side is a linear functional on f . In a steady state, $f(\mathbf{r}, \mathbf{v})$ does not explicitly depend on time but it may still be a function of the spatial position.

Solution of Boltzmann equation is a complicated task even for the simplest scattering models involving scattering of electrons among themselves as well as with phonons and impurities. It is expected from a reasonable model of interaction between the electronic system and the thermal bath that collisions will *restore* the equilibrium distribution function $f_{\text{eq}}(\mathbf{r}, \mathbf{v})$ from any initial distribution – although this property is often difficult to prove mathematically.

† This discussion is from S. Luryi, "Hot electrons in semiconductor devices", in *Hot Electrons in Semiconductors*, ed by N. Balkan, (Oxford University Press, 1997) pp. 385-427. The entire chapter, including all figures and references is available on the internet: Homepage: <http://www.ee.sunysb.edu/~serge/152.dir/152.html>

§ Note that one can also *cool* the carrier system by making it do work against an external field at a fast rate compared to power replenished by the lattice.

Since we know that thermodynamic equilibrium cannot be destroyed by scattering of electrons by a thermal bath, the collision integral must vanish if $f_{\text{eq}}(\mathbf{r}, \mathbf{v})$ is substituted for f . These properties are manifestly expressed in the model which approximates the collision integral by

$$\left[\frac{\partial f}{\partial t} \right]_{\text{coll}} = \frac{f(t, \mathbf{v}) - f_{\text{eq}}(\mathbf{v})}{\tau(\mathbf{v})},$$

where for brevity we have omitted the possible position dependence. The characteristic time constant $\tau(\mathbf{v})$ is called the relaxation time and the whole model is called the relaxation time approximation. In this approximation the perturbed distribution will exponentially relax to equilibrium when the perturbing influence is removed.

The relaxation time approximation is too crude in practice, because different characteristics of the distribution relax with different rates. Thus when collisions are predominantly elastic, it is natural that the first moment $\langle \mathbf{v} \rangle$ of $f(\mathbf{v})$ relaxes rapidly while the second moment $\langle \mathbf{v}^2 \rangle$ takes a long time to relax. The first moment characterizes the electron drift velocity (or average crystal momentum) and the corresponding time constant, τ_m , is called the momentum relaxation time, while the longer time τ_e is called the energy relaxation time. It is often a good approximation to characterize the electron ensemble by a model distribution function, which embodies the relevant relaxation kinetics. The choice of an appropriate distribution function may also depend on time t if the electric fields rapidly vary (on the scale of τ_e or τ_m). Similarly, the choice of a model $f(\mathbf{v})$ may depend on the position \mathbf{r} in the device. Consider hot-electron models, commonly occurring in devices.

Quasi-thermal

An effective temperature T_e is always established in an electronic ensemble of sufficiently high concentration, when the electron-electron (ee) interaction dominates over both phonon and impurity scattering. Starting from any initial distribution, ee collisions lead to an equilibrium within the electron gas. Since the center-of-mass velocity of colliding electrons does not change, the drift velocity $\langle \mathbf{v} \rangle$ of the electron gas remains a constant of the motion. The distribution function is then of the form

$$f(\mathbf{v}) = \left[1 + \exp \left[\frac{m \mathbf{v}^2 / 2 - m \langle \mathbf{v} \rangle \cdot \mathbf{v} - E_F}{kT_e} \right] \right]^{-1}, \quad (\text{F})$$

called the displaced Fermi distribution. It is characterized by the effective temperature T_e , the drift velocity $\langle \mathbf{v} \rangle$, and the Fermi level E_F , determined respectively by the conservation of energy, momentum, and the number of particles. In the reference frame that travels with the velocity $\langle \mathbf{v} \rangle$, this distribution looks like an ordinary Fermi function. For a non-degenerate gas, it reduces to the form of a displaced Maxwellian distribution,

$$f(\mathbf{v}) = e^{-\frac{m \mathbf{v}^2 / 2 - m \langle \mathbf{v} \rangle \cdot \mathbf{v}}{kT_e}}. \quad (\text{M})$$

In the nondegenerate limit $T_e \gg E_F$, the effective temperature T_e is related to the average electron energy $\langle E \rangle$ by the well-known formula,

$$\langle E \rangle = \frac{3}{2} kT_e + \frac{1}{2} m \langle \mathbf{v} \rangle^2.$$

In the opposite limit, $E_F \gg kT_e$, the average energy does not depend on T_e and is determined only by the carrier density n :

$$\langle E \rangle = \frac{3}{5} E_F + \frac{1}{2} m \langle \mathbf{v} \rangle^2, \quad \text{where } E_F = (3\pi^2)^{2/3} \frac{\hbar^2}{2m} n^{2/3}.$$

In the intermediate range, the expression of $\langle E \rangle$ in terms of parameters T_e and E_F can be written in the form of a quadrature.

Any function of the form (F) will make vanish the collision integral that contains only ee interaction. If interaction with impurities is included, the collision integral will vanish only provided $\langle \mathbf{v} \rangle = 0$. Including interaction with phonons, the collision integral will vanish only if T_e coincides with the lattice temperature T .

Overshoot

In the situation when energy is continuously pumped into the electron gas by an electric field or electromagnetic radiation, one cannot ignore other forms of scattering, however slow they may be compared to ee interaction. Indeed, otherwise there would be nothing to check a runaway of the average energy and/or drift velocity. The correct approach is to use the so-called *adiabatic* approximation in which the distribution function is assumed in the form (F) with the effective parameters assumed to follow the variations in the external input, as governed by the balance equations for energy and momentum,

$$\frac{d\langle E \rangle}{dt} = \langle P \rangle - \frac{\langle E \rangle}{\tau_\epsilon} ,$$

$$\frac{d\langle \mathbf{v} \rangle}{dt} = \langle \mathbf{a} \rangle - \frac{\langle \mathbf{v} \rangle}{\tau_m} ,$$

where $\langle P \rangle$ and $\langle \mathbf{a} \rangle$ are, respectively, the average power input and acceleration. For the electronic ensemble (F) moving under a force \mathbf{F} , one has approximately $\langle \mathbf{a} \rangle = \mathbf{F}/m$ and $\langle P \rangle = \mathbf{F} \cdot \langle \mathbf{v} \rangle$. Parameters τ_ϵ and τ_m in these equations are determined self-consistently from the Boltzmann equation – by substituting into the collision integral for slow processes the distribution function in the form (F) – which is assumed maintained by the rapid ee interaction.

Thus, in the adiabatic approximation parameters τ_ϵ and τ_m of the balance equations themselves become functions of the effective temperature T_e . The momentum relaxation time is usually much shorter than the energy relaxation time, $\tau_m \ll \tau_\epsilon$. Indeed, elastic collisions are dominant for momentum relaxation, whereas an effective energy relaxation requires several inelastic interactions with phonons. Therefore, on the time scale of the momentum relaxation, the average electron energy can be considered quasi-static and while $\tau_m = \tau_m(T_e)$ in the balance equation, the effective temperature T_e itself is a function of time, “slowly” rising with a characteristic time τ_ϵ . Thus, at short times ($t < \tau_E$) after the imposition of a strong electric field, the carrier drift occurs with a low-field mobility and the velocity can substantially overshoot its steady-state value. The overshoot phenomenon has become quite important in determining the speed of modern transistors with ultra-short gate lengths. For a 0.25 μm *n*-channel Si MOSFET this hot-electron effect contributes a 20% enhancement in the transistor speed.

The qualitative explanation given above is quite adequate for describing the overshoot in Si and Ge, where there is a strong dependence of the phonon scattering rates on the carrier energy. In GaAs and some other III-V compounds, the dominant scattering process in the Γ valley is due to polar optical phonons, and the rate of these processes at sufficiently high electron energies becomes nearly independent of energy. In such materials the mobility degradation at high energies is associated with the transfer of electrons to the lower-mobility upper valleys. Consequently the overshoot is not seen below the threshold field for the negative differential mobility effect ($F \gtrsim 3.2 \text{ kV/cm}$ in GaAs and $\gtrsim 11 \text{ kV/cm}$ in InP).

Ballistic.

If the force acting on electrons suddenly changes, their subsequent motion for a short time may be considered without taking collisions into account.¹ The time interval $t \lesssim \tau_\beta$ when this is possible, is determined not only by the momentum relaxation time τ_m but also by the characteristic time τ_{ee} of interelectron scattering. By definition, the ee scattering rate is included neither in the rate $1/\tau_m$ nor in the expression $\mu = (e/m) \tau_m$ for steady-state mobility, because ee collisions have no direct effect on $\langle \mathbf{v} \rangle$. Of course, they can have a very strong effect indirectly, by influencing other collision processes through the shape of the distribution function $f(\mathbf{v})$. At carrier concentrations $n \gtrsim 10^{17} \text{ cm}^{-3}$ the characteristic time τ_β for ballistic transport,

$$\frac{1}{\tau_\beta} = \frac{1}{\tau_m} + \frac{1}{\tau_{ee}} ,$$

may be considerably shorter than τ_m .

The concept of ballistic motion also applies to the *steady-state* transport. In this case one considers regions a short distance $d \lesssim \lambda_\beta$ away from an abrupt potential variation. The characteristic length λ_β is related to τ_β by $\lambda_\beta = \langle v \rangle \tau_\beta$ and both quantities depend on the shape of the distribution function $f(\mathbf{v})$. Instead of parameters pertaining to an electron ensemble, one often defines similar parameters for a given state of electron motion, viz. its lifetime $\tau(\mathbf{v})$ and the mean free path $\lambda(\mathbf{v})$, as limited by collisions with the lattice and other electrons. These definitions practically coincide for narrow distributions.

1. The common term “ballistic” applied to this time of motion, conjures up the image of a projectile moving in airless space. This image is not very apt, since interesting properties of the electronic motion in the ballistic regime often depend crucially on the band structure, which defies cannonball analogy.

Steady-state ballistic transport was first demonstrated experimentally in unipolar heterostructures.

In an external field ballistic electrons can be accelerated to velocities much higher than the steady-state saturated velocity v_{sat} [up to the maximum band velocity $dE/d(\hbar\mathbf{k})$]; for electrons in the conduction band of GaAs accelerated in a $\langle 100 \rangle$ direction this limit is $\approx 10^8$ cm/sec] and such an enhancement is important and beneficial for the performance of semiconductor devices. It appears attractive to realize ballistic transport in short-channel field effect transistors, so that the carrier velocity at any point in the channel would be determined by the conservation of energy. Such a situation is realized in vacuum diodes, where the current is space charge limited and is described by the Child-Langmuir law. Despite considerable efforts, no semiconductor device structure has been demonstrated to-date in which the current-voltage characteristics would convincingly conform to a similar law. In practice, conditions for collisionless transport in a 2D channel (where sheet carrier concentrations typically much exceed 10^{11} cm $^{-2}$) are very difficult to realize because of ee collisions (which typically result in $\lambda_{\beta} < 1000 \text{ \AA}$ at these concentrations).

When the rate of ee collisions is very high, as is usually the case in the channel of a field-effect transistor in its *on* state, the electron ensemble behaves rather like a fluid than a gas. In the absence of *other* collisions (with phonons and impurities), the electronic fluid in the FET channel is described by hydrodynamic equations similar to those for shallow water (Dyakonov and Shur, 1993). Based on this analogy, Dyakonov and Shur discussed several new effects related to plasma oscillations in the 2D electron fluid. In particular, a short-channel high-mobility transistor has a resonant response to an electromagnetic radiation at the plasma wave frequency of the 2D electrons (Dyakonov and Shur, 1996). This effect can be used to implement detectors, mixers and multipliers at terahertz frequencies. As pointed out by the authors, these devices should operate at much higher frequencies than conventional, transit-time limited devices, since the plasma waves propagate much faster than electrons. Moreover, their responsivities and conversion efficiencies can be expected to greatly exceed those of Schottky diodes currently used as detectors, mixers and multipliers in the terahertz range.

Mesoscopic.

When the carrier concentration is sufficiently low and ee collisions are rare, the shape of their distribution function under external perturbation may depart considerably from the quasi-equilibrium form (F). The Maxwellian shape (M) (Gaussian in velocities) can be viewed – in the spirit of the well-known central limit theorem of statistics – as resulting from large number of independent scattering events, each contributing or withdrawing a random amount of energy.† Besides ee collisions, the electron distribution function can be maxwellized by other interactions as well, provided independent scattering events exchange random energies with electrons. Scattering by acoustic phonons has this property, while optical phonon scattering does not. If the latter were the only inelastic interaction, the equilibrium shape of the electron distribution function would be rather strange and thermodynamic properties of the electron ensemble rather different. The peculiarity of interaction with optic phonons stems from their largely monochromatic nature, which quantizes the energy exchange in units of HO .

At sufficiently high temperatures, the electron energy relaxation rate due to optical phonons ($1/\tau^{(op)}$) is higher than that due to acoustic phonons ($1/\tau^{(ac)}$) by several orders of magnitude. Typically, in semiconductors $\tau^{(op)} \lesssim 10^{-12}$ s and $\tau^{(ac)} \gtrsim 10^{-9}$ s (Conwell, 1967). This disparity of the inelastic relaxation times can lead to the formation of an electronic ensemble that is in equilibrium with the optical-phonon system but has not yet appreciably interacted with acoustic phonons. Manifestation of these properties in electronic transport can be conveniently referred to as the “classical mesoscopic effects” (drawing a parallel to and a distinction from the quantum mesoscopic effects that occur when the coherence length or time of an electronic wave function exceeds characteristic system dimensions. Grinberg and Luryi (1990) considered the kinetics of an electron ensemble initially characterized by a Maxwellian distribution with $T_e = T_l$, subject to interaction with the lattice at equilibrium temperature T . Since the initial distribution is not the equilibrium Boltzmann function, it evolves in time because of the electron interactions with optical phonons, acoustic phonons, and due to the e-e scattering. Under the

† The form of distribution (F) does not contradict the central limit theorem. In the presence of the quantum correlation between electrons, maintained by the Pauli exclusion principle, multiple collisions lead to a Fermi rather than Maxwell distribution.

assumption that $\tau^{(\text{op})}/\tau^{(\text{ac})} \ll 1$ and $\tau^{(\text{op})}/\tau_{ee} \ll 1$, electrons rapidly establish a quasi equilibrium with the optical phonon field (the mesoscopic state) and then – on a longer scale – the true equilibrium is established by other inelastic scattering processes. Even though the e-e scattering does not change the average electron energy, it counts as an inelastic interaction, because it changes the shape of the non-stationary distribution.

The distribution function of the mesoscopic state can be determined from the statistical consideration alone, without actually solving the kinetic equation. The shape of this function is illustrated in Fig. 10 both for the narrow ($T_i < T$) and broad ($T_i > T$) initial distributions. Even though the electron system is in perfect equilibrium with optical phonons at temperature T , thermodynamic properties of electrons in the mesoscopic state are very different from those in true equilibrium, e.g., the average energy $\langle E \rangle \neq (3/2) kT$ and the specific heat deviates from the classical value $3k/2$ (Grinberg and Luryi, 1990). Also the electron mobility in the mesoscopic state shows a strong overshoot. The time scale of this effect is very different from the conventional velocity overshoot; the mesoscopic state is established in less than 1 ps and persists for a long time (up to nanoseconds!) – controlled by acoustic phonon and interelectron scattering.

